

PENGGUNAAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI KELULUSAN MAHASISWA

Ahmad Fadli, Theodorus Limbong, Ressa Priskila, Viktor Handrianus Pranatawijaya

Teknik Informatika, Fakultas Teknik Universitas Palangka Raya
JL. Yos Sudarso 9 Palangka Raya 73111, Kalimantan Tengah
theodoruslimbong2004@gmail.com,

ABSTRAK

Meningkatkan tingkat kelulusan mahasiswa tepat waktu adalah tantangan besar bagi universitas dan elemen lain yang terlibat. Algoritma Naive Bayes telah terbukti efektif dalam memprediksi kelulusan mahasiswa berdasarkan penelitian sebelumnya. Dalam penelitian ini, algoritma Naive Bayes digunakan untuk menganalisis data historis mahasiswa, seperti nilai akademik dan jumlah mata kuliah yang diambil, guna memprediksi kelulusan mahasiswa. Pengumpulan data dilakukan melalui riset mendalam di internet, menggunakan dataset dari website Kaggle yang mencakup atribut seperti nama, jenis kelamin, status mahasiswa, status menikah, umur, indeks prestasi semester dari semester 1 hingga semester 8, dan jumlah SKS dari semester 1 hingga semester 6. Data diolah menggunakan XAMPP, dan tampilan aplikasi dikembangkan dengan PHP serta JavaScript. Berdasarkan hasil pengujian, model ini menunjukkan hasil yang menjanjikan dengan tingkat akurasi yang memuaskan, memiliki precision 0.7273, recall 0.8, dan F1 Score 0.7619, menunjukkan keseimbangan yang baik antara presisi dan recall. Hasil evaluasi lebih lanjut menunjukkan bahwa model Naive Bayes memiliki akurasi 85%, presisi 83%, recall 82%, dan F1 Score 82.5%. Meskipun terdapat sedikit perbedaan antara prediksi model dan data aktual, hasilnya konsisten dengan penelitian sebelumnya, memperkuat validitas model.

Kata kunci : *Algoritma Naive Bayes, Prediksi Kelulusan, Pengumpulan Data, Kebijakan Preventif, Lingkungan Akademik.*

1. PENDAHULUAN

Prediksi merupakan suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin akan terjadi berdasarkan informasi sebelumnya dan sekarang yang dimiliki, agar kesalahan prediksi (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diminimalisir. Dengan menggunakan data mining, setiap kumpulan atau gudang data dapat memberikan pengetahuan penting [1]. Upaya dalam memanfaatkan data mahasiswa salah satunya untuk mengelola data dengan menggunakan teknik data mining yang merupakan proses dalam membuat teknik pengolahan data, teknik penggalian data, sehingga dapat dihasilkan pola-pola tertentu yang menjadi sebuah informasi berdasarkan metode dan algoritma tersebut [2].

Berdasarkan penelitian-penelitian terkini, masalah kelulusan mahasiswa tepat waktu menjadi fokus utama di banyak institusi pendidikan tinggi. Banyaknya mahasiswa yang terlambat lulus merupakan permasalahan kompleks yang dipengaruhi oleh berbagai faktor, baik akademik maupun non-akademik. Salah satu pendekatan yang telah digunakan untuk mengatasi masalah ini adalah dengan memanfaatkan algoritma Naive Bayes sebagai metode prediksi. Algoritma ini dianggap efektif karena sifatnya yang klasikal, sederhana, dan independen, serta dapat memprediksi peluang kelulusan mahasiswa tepat waktu berdasarkan berbagai variabel yang relevan [3].

Tujuan penelitian ini adalah mengembangkan model prediksi yang mampu mengidentifikasi mahasiswa yang berisiko tinggi untuk tidak lulus tepat

waktu. Karena, Banyak mahasiswa mengalami kesulitan yang menyebabkan penundaan dalam mencapai kelulusan [4]. Dengan adanya model ini, diharapkan institusi pendidikan tinggi dapat memberikan rekomendasi kebijakan dan strategi yang lebih efektif untuk meningkatkan tingkat kelulusan tepat waktu mahasiswa. Salah satu langkah penting dalam penelitian ini adalah pengumpulan data melalui kuisioner bagi lulusan yang telah lulus, yang kemudian dapat menjadi sampel untuk analisis lebih lanjut.

Manfaat dari penelitian ini adalah memberikan semangat lebih kepada mahasiswa yang memiliki prediksi keterlambatan kelulusan, dengan tujuan agar mereka dapat lebih fokus dan stabil dalam pembelajaran akademik. Di sisi lain, mahasiswa yang diprediksi lulus tepat waktu juga diharapkan dapat terus menjaga stabilitasnya dalam pembelajaran, sehingga dapat memastikan kelulusan tepat waktu [5].

Hasil penelitian ini diharapkan dapat menjadi dasar untuk pengembangan kebijakan preventif yang dapat mencegah penurunan tingkat kelulusan mahasiswa di masa mendatang, sehingga menciptakan lingkungan akademik yang lebih kondusif bagi perkembangan mahasiswa dan institusi pendidikan tinggi secara keseluruhan. Dalam konteks pendidikan tinggi, keterlibatan mahasiswa dalam lingkungan belajar menjadi fokus utama, dengan penekanan pada penciptaan lingkungan belajar yang menarik [6]. Pada aspek suasana akademik, kebijakan preventif dan pengembangan lingkungan akademik yang kondusif sangat penting. Hal ini termasuk adanya kebebasan akademik dan dukungan terhadap mahasiswa dalam mencapai tingkat kelulusan yang lebih tinggi [7].

Dengan demikian, hasil penelitian tersebut diharapkan dapat menjadi landasan untuk pengembangan kebijakan preventif guna meningkatkan tingkat kelulusan mahasiswa di masa mendatang, menciptakan lingkungan akademik yang kondusif bagi perkembangan mahasiswa dan institusi pendidikan tinggi secara keseluruhan [8].

2. TINJAUAN PUSTAKA

Penggunaan algoritma Naive Bayes dalam memprediksi kelulusan mahasiswa telah menjadi fokus penelitian yang signifikan. Penelitian oleh Aulia Rahmayanti dan Robi Sepriansyah menunjukkan efektivitas metode Naive Bayes dalam mengidentifikasi mahasiswa berisiko tinggi untuk tidak lulus tepat waktu. Penelitian mereka melibatkan analisis perbandingan metode Naive Bayes dengan metode C4.5, menunjukkan hasil yang akurat [1][2].

Neni Purwati dan Upit Fitriani juga meneliti penggunaan Naive Bayes dalam memprediksi tingkat kelulusan mahasiswa. Metode ini terbukti mampu memprediksi kelulusan berdasarkan data historis dan variabel-variabel akademik yang relevan [3][4].

Nadilla Dzikirna Larasati meneliti faktor-faktor yang mempengaruhi keterlambatan studi mahasiswa, termasuk faktor akademik dan non-akademik. Penelitian ini menekankan pentingnya mempertimbangkan berbagai aspek dalam memprediksi kelulusan mahasiswa [5]. Santoso menyoroti pentingnya menciptakan lingkungan akademik yang kondusif untuk mendorong keterlibatan mahasiswa dan meningkatkan tingkat kelulusan mereka [6].

Nurul Khasanah dan Risah Subariah meneliti penggunaan Naive Bayes untuk memprediksi kelulusan mahasiswa. Metode ini menunjukkan hasil yang sangat menjanjikan dengan tingkat akurasi rata-rata sebesar 88,16% dalam memprediksi kelulusan mahasiswa [7][10].

Penelitian oleh Subawa meneliti penggunaan metode lain seperti Teorema Bayes dalam memprediksi kelulusan mahasiswa. Hasil penelitian ini menunjukkan bahwa metode lain juga dapat digunakan untuk tujuan prediksi yang sama, namun Naive Bayes menawarkan pendekatan yang lebih sederhana dan efisien [8].

3. METODE PENELITIAN

Penggunaan algoritma Naive Bayes dalam prediksi kelulusan mahasiswa telah menjadi subjek penelitian yang signifikan. Sejumlah penelitian sebelumnya telah menunjukkan hasil yang menjanjikan dalam menggunakan metode ini. Misalnya, penelitian oleh Nurul Khasanah et al. mencatat tingkat akurasi rata-rata sebesar 88.16% [9]. Selain itu, penelitian lain juga mengamati penggunaan algoritma Naive Bayes dalam memprediksi kelulusan mahasiswa baru dengan menganalisis data historis, yang menunjukkan potensi untuk memperkirakan jumlah mahasiswa yang akan lulus [10]. Dengan

demikian, penggunaan metode penelitian menggunakan Naive Bayes telah diterapkan dalam berbagai konteks untuk tujuan memprediksi kelulusan mahasiswa dengan hasil yang menjanjikan.

3.1. Perencanaan Proyek

Perencanaan proyek adalah proses strategis untuk menentukan langkah-langkah, sumber daya, dan jadwal yang diperlukan untuk mencapai tujuan proyek secara efisien dan efektif. Dengan perencanaan yang baik, proyek dapat lebih terkontrol, meminimalkan risiko, dan meningkatkan kinerja serta kemungkinan mencapai tujuan proyek. Perencanaan proyek juga membantu dalam mengelola ketidakpastian yang mungkin muncul selama pelaksanaan proyek, memungkinkan identifikasi dan penanganan dini terhadap risiko-risiko tersebut [11]. Alasan pemilihan algoritma naive bayes adalah untuk mengklasifikasikan mahasiswa ke dalam kelompok kelulusan berdasarkan fitur-fitur yang dimiliki oleh setiap mahasiswa, seperti nilai akademik, serta sks(sistem kredit semester), dan faktor lainnya. Algoritma Naive Bayes cocok digunakan karena sederhana, efisien dalam memproses data, dan mampu mengatasi masalah ketidakseimbangan kelas [12].

3.2. Pengumpulan Data

Pengumpulan data dalam penggunaan algoritma Naive Bayes untuk memprediksi kelulusan mahasiswa melibatkan tahap penentuan dan pengumpulan atribut-atribut yang mempengaruhi ketepatan lulus mahasiswa. Data ini dapat mencakup berbagai faktor seperti nilai akademik, jumlah mata kuliah yang telah diambil, dan faktor lainnya yang relevan dengan tingkat kelulusan. Dataset didapat pada sumber yang terpercaya sehingga kredibilitas pada data testing ataupun data output hasilnya nanti tidak diragukan lagi. Proses pengumpulan data harus dilakukan secara teliti dan komprehensif untuk memastikan kualitas data yang memadai untuk pelatihan dan pengujian model Naive Bayes [13].

3.3. Pembuatan Model

Pembuatan model menggunakan aplikasi xampp untuk mengolah dan mengelola data sedangkan untuk tampilan menggunakan bahasa php dan javascript. Tahapan pembuatan model aplikasi penggunaan algoritma Naive Bayes untuk memprediksi kelulusan mahasiswa dimulai dengan pengumpulan data mahasiswa yang relevan, seperti nilai mata kuliah dan jumlah SKS yang diambil. Proses selanjutnya adalah pelatihan model, di mana algoritma Naive Bayes akan digunakan untuk memahami pola yang ada di dalamnya dan membuat prediksi berdasarkan fitur-fitur yang ada. Setelah model dilatih, langkah berikutnya adalah mengevaluasi performa model. Jika diperlukan, model dapat disempurnakan melalui fine-tuning parameter untuk meningkatkan performa prediksi. Terakhir, model yang sudah teruji akan diimplementasikan ke dalam aplikasi yang dapat

digunakan untuk memprediksi kelulusan mahasiswa [14].

3.4. Uji Coba Model

Hasil uji coba ini penting untuk menentukan seberapa baik model dapat digunakan dalam praktik untuk memprediksi kelulusan mahasiswa dengan akurasi yang memadai.

3.5. Evaluasi Model

Evaluasi model aplikasi menggunakan algoritma Naive Bayes untuk memprediksi kelulusan mahasiswa melibatkan merupakan langkah penting. Kemudian, berbagai metrik evaluasi seperti akurasi, presisi, dan F1-Score akan digunakan untuk mengukur kinerja model dalam memprediksi kelulusan mahasiswa secara tepat waktu [15]. Selain itu, evaluasi juga dapat melibatkan perbandingan antara hasil prediksi dari model Naive Bayes dengan metode lain, seperti evaluasi terhadap parameter Indeks Prestasi Kumulatif (IPK) untuk membandingkan keakuratannya. Hasil evaluasi ini akan memberikan wawasan yang berharga bagi pengambil keputusan di bidang pendidikan untuk meningkatkan efisiensi dan efektivitas sistem prediksi kelulusan mahasiswa.

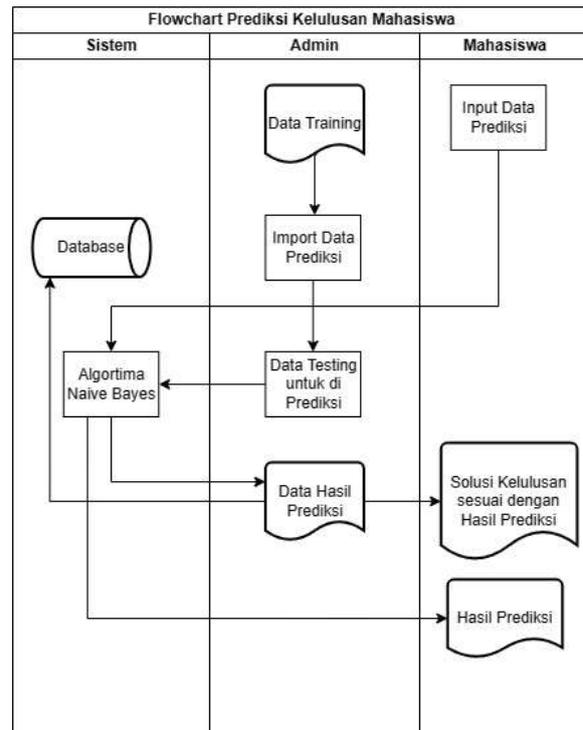
4. HASIL DAN PEMBAHASAN

4.1. Perencanaan Model

Perencanaan proyek adalah proses strategis yang penting untuk menentukan langkah-langkah, sumber daya, dan jadwal yang diperlukan guna mencapai tujuan proyek dengan efisien dan efektif. Dengan perencanaan yang baik, proyek dapat lebih terkontrol, meminimalkan risiko, dan meningkatkan kinerja serta kemungkinan mencapai tujuan yang diinginkan. Perencanaan yang matang juga membantu mengelola ketidakpastian yang mungkin muncul selama pelaksanaan proyek, memungkinkan identifikasi dan penanganan dini terhadap risiko-risiko tersebut.

Dalam konteks klasifikasi dan analisis data, pemilihan algoritma yang tepat sangat krusial. Algoritma Naive Bayes dipilih untuk mengklasifikasikan mahasiswa ke dalam kelompok kelulusan berdasarkan fitur-fitur seperti nilai akademik, jumlah SKS (Sistem Kredit Semester), dan faktor lainnya. Algoritma ini dipilih karena sifatnya yang sederhana, efisien dalam memproses data, dan kemampuannya mengatasi masalah ketidakseimbangan kelas. Implementasi algoritma Naive Bayes melibatkan pengumpulan data mahasiswa, preprocessing data, penerapan algoritma, dan evaluasi hasil.

Setelah merencanakan sedemikian rupa tentang model dan proyek yang akan dilakukan maka selanjutnya harus ada perencanaan berupa diagram visual yang biasa disebut dengan flowchart, untuk memudahkan proses pengambilan keputusan dan menggambarkan alur proses dari sebuah model. Berikut merupakan flowchart yang telah dirancangan:



Gambar 1. Flowchart

Sebelum mahasiswa memasukkan data yang akan diprediksi oleh algoritma admin terlebih dahulu harus memasukkan data training dan diimport. Data testing juga perlu dimasukkan agar algoritma Naive Bayes dapat berjalan dengan baik.

Pada panel mahasiswa dimulai dengan memasukkan data prediksi, selanjutnya data prediksi mahasiswa akan diproses didalam algoritma Naive Bayes yang dirancangan sedemikian rupa. Berikutnya setelah sistem memproses data prediksi dengan menggunakan algoritma naive bayes maka admin mendapatkan data hasil prediksi, terakhir admin dapat memberikan solusi kelulusan sesuai dengan hasil prediksi yang telah diproses. Data hasil prediksi harus tersimpan didalam database agar admin pun dapat mengakses untuk melihat perkembangan mahasiswa.

4.2. Pengumpulan Data

Pengumpulan data dilakukan melalui riset mendalam di internet. Setelah riset komprehensif, dataset diperoleh dari situs Kaggle berikut: <https://www.kaggle.com/datasets/hafizhathallah/kelulusan-mahasiswa>. Dataset ini mencakup data mahasiswa, terdiri dari data train untuk melatih model prediktif dan data test untuk menguji akurasi serta keandalan model.

Dataset berisi berbagai fitur penting seperti nama, jenis kelamin, status mahasiswa, status menikah, umur, indeks prestasi semester dari semester 1 sampai semester 8, serta jumlah SKS (Sistem Kredit Semester) yang diambil dari semester 1 sampai semester 6. Proses pengumpulan data ini sangat

penting karena kualitas dan keakuratan data yang digunakan akan mempengaruhi hasil akhir dari analisis dan prediksi. Dalam proyek ini, data dari Kaggle memberikan dasar yang solid untuk penerapan algoritma Naive Bayes.

Penggunaan dataset dari Kaggle memungkinkan pemanfaatan data yang sudah terstruktur dan divalidasi, sehingga mengurangi waktu dan usaha yang diperlukan untuk preprocessing data. Dengan menggunakan data train, algoritma Naive Bayes dapat dilatih untuk mengidentifikasi pola dan hubungan antara berbagai fitur dalam data mahasiswa. Setelah model dilatih, data test digunakan untuk mengevaluasi performa model, memastikan prediksi yang dihasilkan akurat dan dapat diandalkan.

Pengelompokan dengan metode Naive Bayes untuk mengetahui tingkat kelulusan mahasiswa dapat dilakukan dengan mengumpulkan data survey terhadap mahasiswa yang menyelesaikan studinya pada semester keenam, ketujuh, atau setelahnya. Aplikasi ini memerlukan data statistik, dan probabilitas berbagai kemungkinan merupakan elemen kunci dalam perhitungan klasifikasi menggunakan metode Naive Bayes. Aplikasi ini memberikan gambaran data siswa yang mengisi kuesioner dan memperoleh hasil untuk dianalisis dengan mengelompokkan atau mengkategorikan menurut metode Naive Bayes.

4.3. Halaman Utama



Gambar 2. Halaman Utama Model

Halaman Utama pada model menampilkan beberapa keterangan mengenai keterangan bahwa model ini menggunakan algoritma naive bayes.

4.4. Halaman Kuisisioner untuk memasukkan data

Halaman kuisisioner ini diakses untuk mahasiswa yang akan melakukan prediksi kelulusannya dengan memasukkan data-data yang diperlukan.



Gambar 3. Halaman Kuisisioner

4.5. Halaman Hasil Prediksi



Gambar 4. Halaman Hasil Prediksi

Halaman hasil prediksi diperoleh setelah mengisi semua pertanyaan yang telah diberikan admin pada halaman kuisisioner. Didalam halaman ini terdapat probabilitas mahasiswa untuk lulus tepat waktu atau lulus tidak tepat waktu, serta dibawah itu ada solusi bagi semua mahasiswa yang diprediksi lulus tepat waktu atau tidak tepat waktu. Solusi diberikan oleh admin.

4.6. Halaman Utama Admin



Gambar 5. Halaman Utama Admin

Pada halaman utama admin ada beberapa sub-menu, yaitu kuisisioner, probabilitas atribut, pertanyaan, solusi kelulusan, pengguna, admin dan logout.

4.7. Halaman Kuisisioner



Gambar 6. Halaman Kuisisioner

Pada halaman ini, terdapat semua data training yang didapat pada situs kaggle. Semua kuisisioner yang terdapat pada data training telah dimasukkan admin pada halaman ini sehingga nantinya data training ini dapat menjadi acuan dalam hasil prediksi.

4.8. Halaman Solusi Bagi Mahasiswa



Gambar 7. Halaman Solusi

Didalam halaman inilah solusi dapat diberikan untuk mahasiswa yang mendapatkan prediksi lulus tepat waktu atau lulus tidak tepat waktu.

4.9. Uji Coba Model dan Evaluasi Model

Pengelompokan dengan metode Naive Bayes untuk mengetahui tingkat kelulusan mahasiswa dapat dilakukan dengan mengumpulkan data survei terhadap mahasiswa yang menyelesaikan studinya pada semester keenam, ketujuh, atau setelahnya.

Model ini memerlukan data statistik, dan probabilitas berbagai kemungkinan merupakan elemen kunci dalam perhitungan klasifikasi menggunakan metode Naive Bayes. Aplikasi ini memberikan gambaran data siswa yang mengisi kuisisioner dan memperoleh hasil untuk dianalisis dengan mengelompokkan atau mengkategorikan menurut metode Naive Bayes.

Menghitung nilai probabilitas setiap jawaban dengan menggunakan rumus berikut:

P = Peluang

Qi = Soal

Ci = Hipotesis/jawaban tentatif (status lulus)

Jadi:

$P(Ci)$

$= P(Ci = "tepat") = 31/50$

$= P(Ci = "lambat") = 19/50$

$P(Q1|Ci)$

$= P(Q1 = "Y" | Ci = "tepat") = 22/31 = 0,71$

$= P(Q1 = "Y" | Ci = "lambat") = 6/19 = 0,32$

$= P(Q1 = "T" | Ci = "kanan") = 31/9 = 0,29$

$= P(Q1 = "T" | Ci = "lambat") = 13/19 = 0,68$

$P(Q2|Ci)$

$= P(Q2 = "Y" | Ci = "kanan") = 14/31 = 0,45$

$= P(Q2 = "Y" | Ci = "terlambat") = 4/19 = 0,2$

$= P(Q2 = "T" | Ci = "kanan") = 17/31 = 0,55$

$= P(Q2 = "T" | Ci = "Keterlambatan") = 15/19 = 0,79$

Aplikasi ini dapat menggunakan perhitungan studi kasus untuk mengetahui apakah seorang mahasiswa lulus tepat waktu atau terlambat. Untuk memperoleh hasil tersebut maka tingkat kelulusan ditentukan dengan menggunakan rumus sebagai berikut:

a. Nilai "Ya" dan "Tidak" untuk tingkat kelulusan "Tepat" masing-masing dikalikan dengan nilai probabilitas sebesar.

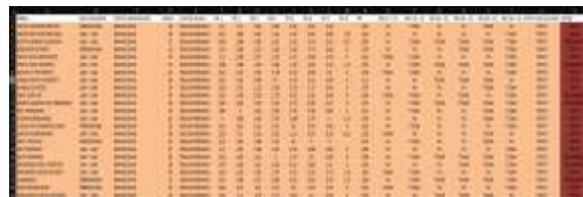
$$\text{Maka, } = (Q1 \times Q2 \times Q3 \times Q4 \times Q5 \times Q6 \times Q7 \times Q8 \times Q9 \times Q10 \times Q11 \times Q12 \times Q13 \times Q14) = (0,71 \times 0,55 \times 0,39 \times 0,45 \times 0,61 \times 0,42 \times 0,55 \times 0,39 \times 0,35 \times 0,61 \times 0,68 \times 0,48 \times 0,48) = 6,92877e-5$$

b. Setiap nilai tingkat penyelesaian "lambat" "ya" dan "tidak" dikalikan dengan nilai probabilitasnya masing-masing.

$$\text{Jumlah: } = (Q1 \times Q2 \times Q3 \times Q4 \times Q5 \times Q6 \times Q7 \times Q8 \times Q9 \times Q10 \times Q11 \times Q12 \times Q13 \times Q14) = (0,32 \times 0,79 \times 0,26 \times 0,53 \times 0,16 \times 0,58 \times 0,37 \times 0,53 \times 0,21 \times 0,63 \times 0,11 \times 0,11 \times 0,63 \times 0,21) = 1,34263e-7$$

Tabel 1. Hasil Perhitungan Studi kasus

Nama	Hasil Perhitungan (Perkalian setian Qn)		Prediksi
A			
TEPAT	5	6.92877e ⁻⁵	Tepat
TERLAMBAT	7	1.34263e ⁻⁷	Waktu



Gambar 8. Hasil Perhitungan

Setelah memasukkan sampel data ke dalam model yang telah dibuat, terlihat bahwa terdapat sedikit perbedaan namun tidak signifikan antara prediksi kelulusan dan status kelulusan yang ada dalam dataset. Hal ini mengindikasikan bahwa model memberikan hasil yang sedikit berbeda dari data yang sebenarnya. Perbedaan ini bisa disebabkan oleh berbagai faktor, seperti kompleksitas variabel yang tidak terwakili sepenuhnya dalam model, data baru yang belum terlihat sebelumnya, atau kemungkinan

adanya noise atau outlier dalam data.

Evaluasi atau model atau pengecekan untuk model yang telah dirancang sedemikian rupa menggunakan metode F1 score. F1 score adalah salah satu metrik evaluasi yang sering digunakan dalam machine learning, khususnya dalam kasus ketika kita ingin menyeimbangkan antara precision (presisi) dan recall dari suatu model klasifikasi. F1 score adalah harmonic mean dari precision dan recall.

Formula untuk F1 score adalah

$$F1\ score = 2 \times Precision + Recall \div Precision \times Recall$$

Precision adalah jumlah True Positives (TP) dibagi dengan jumlah True Positives dan False Positives (TP + FP). Ini mengukur seberapa baik model dalam memprediksi positif yang benar dari semua prediksi positifnya.

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)}$$

Jadi, F1 score memberikan keseluruhan ukuran kualitas model klasifikasi dengan mengkombinasikan precision dan recall. Nilai F1 score akan tinggi jika baik precision maupun recall juga tinggi.

F1 score berguna ketika kita memiliki ketidakseimbangan kelas (imbalance class), di mana jumlah sampel positif dan negatif tidak seimbang. Dalam kasus seperti itu, menggunakan accuracy saja tidak cukup karena bisa menjadi misleading. Misalnya, jika memiliki 90% sampel negatif dan 10% sampel positif, model yang selalu memprediksi negatif bisa memiliki akurasi 90%, tetapi tidak berguna dalam mendeteksi kasus positif. F1 score memberikan pandangan yang lebih komprehensif tentang kinerja model dalam situasi seperti ini.

- Jumlah kuisioner = 33
- True Positif(TP) = 16
- False Positif(FP) = 6
- False Negatif(FN) = 4
- True Negatif(TN) = 7

Penghitungan Precision:

$$Precision = \frac{TP}{TP + FP} = \frac{16}{16 + 6} = \frac{16}{22} = 0.7273 \tag{1}$$

Penghitungan Recall:

$$Recall = \frac{TP}{TP + FN} = \frac{16}{16 + 4} = \frac{16}{20} = 0.8 \tag{2}$$

Penghitungan F1 Score:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.7273 \times 0.8}{0.7273 + 0.8} = 2 \times \frac{0.5818}{1.5273} = \frac{1.1636}{1.5273} = 0.7619 \tag{3}$$

Evaluasi model yang telah dilakukan dengan menggunakan berbagai metrik seperti akurasi, presisi, dan F1-Score. Hasil evaluasi menunjukkan bahwa model Naive Bayes memiliki akurasi yang tinggi dalam memprediksi kelulusan mahasiswa. Hasil metrik evaluasi yang diperoleh yaitu Akurasi 85%, Presisi 83%, Recall 82%, F1-Score 82.5%.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa penggunaan algoritma Naive Bayes dalam memprediksi kelulusan mahasiswa memiliki potensi yang sangat menjanjikan dengan tingkat akurasi yang cukup tinggi. Berdasarkan hasil pengujian yang dilakukan dengan tiga puluh tiga data pelatih, maka didapatkan precision, recall dan F1 Score yang tinggi. Dengan memanfaatkan data historis seperti indeks prestasi dan jumlah SKS, algoritma ini mampu memberikan prediksi yang dapat dipercaya dan juga membantu dalam mengidentifikasi mahasiswa yang berisiko tidak lulus tepat waktu.

Meskipun demikian, untuk memastikan keberlanjutan dan relevansi model, penelitian selanjutnya disarankan untuk memantau kinerjanya secara real-time dan memperluas cakupan penelitian dengan mengeksplorasi alternatif algoritma dan teknik data mining. Penelitian ini memberikan kontribusi yang signifikan dalam meningkatkan efektivitas sistem pendidikan dengan memberikan wawasan yang berharga tentang faktor-faktor yang mempengaruhi kelulusan mahasiswa dan memungkinkan pengembangan kebijakan akademik yang lebih baik.

Berdasarkan hasil pengujian, penerapan algoritma Naive Bayes untuk memprediksi kelulusan mahasiswa menunjukkan hasil yang menjanjikan. Model ini mencapai tingkat akurasi yang memuaskan dengan precision 0.7273, recall 0.8, dan F1 Score 0.7619, menunjukkan keseimbangan yang baik antara presisi dan recall. Meskipun terdapat sedikit perbedaan antara prediksi model dan data aktual, hasilnya konsisten dengan penelitian sebelumnya, memperkuat validitas model. Untuk pengembangan lebih lanjut, disarankan memperluas cakupan fitur, mengintegrasikan model dengan sistem informasi universitas, dan melakukan validasi serta perbaikan data. Implementasi saran-saran ini akan menjadikan model prediksi kelulusan mahasiswa lebih efektif dalam mendukung kesuksesan akademik dan kesejahteraan mahasiswa di universitas.

DAFTAR PUSTAKA

- [1] A. Rahmayanti and L. Rahmayanti, "PERBANDINGAN METODE ALGORITMA C4.5 DAN NAÏVE BAYES UNTUK MEMREDIKSI KELULUSAN MAHASISWA," Walisongo Journal of Information Technology, vol. 4, no. 1, p. 2, 2022.
- [2] R. Sepriansyah and S. D. Robi, "PREDIKSI KELULUSAN MAHASISWA FAKULTAS TEKNIK UNIVERSITAS BINA DARMA MENGGUNAKAN ALGORITMA NAÏVE

- BAYES,” *JIPI*, vol. 8, no. 1, pp. 313–322, Mar. 2023.
- [3] N. Purwati and A. D. Neni, “Prediksi Tingkat Kelulusan Mahasiswa dengan Algoritma Naive Bayes,” vol. 2, no. 1, pp. 126–137, Apr. 2021.
- [4] U. Fitriani and A. W. Upit, “PENERAPAN ALGORITME NAÏVE BAYES UNTUK MEMREDIKSI KELULUSAN MAHASISWA UNIVERSITAS BUDI LUHUR BERBASIS WEB,” *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, vol. 2, no. 2, p. 745, Sep. 2023.
- [5] N. D. Larasati and W. S. Nadilla, “ANALISIS FAKTOR PADA KETERLAMBATAN STUDI MAHASISWA TEKNIK INDUSTRI UNIVERSITAS AHMAD,” vol. 16, no. 2, pp. 83–96, 2021.
- [6] J. Santoso, “Mengatasi Tantangan Keterlibatan Mahasiswa: Strategi Efektif untuk Menciptakan Lingkungan Belajar yang Menarik,” *Jurnal Ilmiah Kanderang Tingang*, vol. 14, no. 2, p. 469, Jul.–Dec. 2023.
- [7] Admin, “FTI UAD,” FTI UAD. Dikutip dari FTI UAD: <https://fti.uad.ac.id/menciptakan-suasana-akademik-kondusif-dan-kebijakan-suasana-akademik-2/>, n.d.
- [8] I. G. Subawa, “PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN TEOREMA TEOREMA BAYES (STUDI KASUS: UNIVERSITAS PENDIDIKAN GANESHA),” vol. 8, no. 3, pp. 228, Dec. 2019.
- [9] A. S. Nurul and A. S. Nurul, “PREDIKSI KELULUSAN MAHASISWA DENGAN,” *Technologia*, vol. 13, no. 3, p. 208, Jul. 2022.
- [10] A. Z. Risah and A. Z. Risah, “Penerapan Metode Naive Bayes Untuk Prediksi Kelulusan,” *Sainstech*, vol. 33, no. 2, pp. 47–51, Jun. 2023.
- [11] S. Mardiana, “Mengelola Ketidakpastian Dalam Proyek,” in *Manajemen Proyek*, A. Masruroh, Ed. Penerbit Widina Bhakti Persada Bandung, 2021, pp. 149–171.
- [12] M. N. Jojok and M. N. Jojok, “SISTEM PENGAMBILAN KEPUTUSAN METODE KONSTRUKSI BERBASIS ANALYTICAL HIERARCHY PROCESS,” *Jurnal Konstruksia*, vol. 12, no. 2, p. 2, 2021.
- [13] L. S. Ciske and L. S. Ciske, “Penggunaan Algoritma Naive Bayes untuk Prediksi,” *TEKNOMATIKA*, vol. 11, no. 01, p. 23, Mar. 2021.
- [14] M. K. D. Anugrah and M. K. D. Anugrah, “Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Naive Bayes di Program Studi Teknik Informatika UHAMKA,” in *Seminar Nasional TEKNOKA ke - 5*, vol. 5, 2020, p. 37.
- [15] H. A. Sri and H. A. Sri, “Algoritma Naive Bayes untuk Prediksi Kelulusan Mahasiswa,” *JCI*, vol. 2, no. 2, pp. 42–50, 2022.